2008

Using Spatial Analyses to Examine Student Proficiency: Guiding District Consolidation and Reform Policy Decisions

Monica A. Haddad
Iowa State University, haddad@iastate.edu

Thomas Alsbury
North Carolina State University at Raleigh

Abstract

Nationally, all districts, regardless of their size, face the challenge to ensure that all students have access to a quality education. However, some characterize the alternatives for rural families as being between schools that are too small or too far away, and for urban families as being between schools that are too large or too far away. State policies to equalize educational spending, cap bond and levy rates, require common teacher salary schedules, and consolidate school districts are geared toward improving quality and attempting to resolve problems of inefficiency. However, despite the use of the state policy to equalize educational quality, educational gaps persist. The most identifiable achievement differences between and among school districts are well documented and result from variations in (a) teacher quality; (b) teacher mobility; (c) student characteristics, like poverty, ethnicity, gender, and mobility; and (d) community wealth (Goldhaber & Anthony, 2004; Haselton, 2004; Parcel & Dufur, 2004; Uribe, 2003, 2004).

Using Spatial Analyses to Examine Student Proficiency: Guiding District Consolidation and Reform Policy Decisions

Article · January 2008

2 authors:


Monica A. Haddad
Iowa State University


Thomas Alsbury
Northwest University

Planning and Changing
Vol. 39, No. 1&2, 2008, pp. 98–126

USING SPATIAL ANALYSES TO EXAMINE STUDENT PROFICIENCY: GUIDING DISTRICT CONSOLIDATION AND REFORM POLICY DECISIONS

Introduction

Nationally, all districts, regardless of their size, face the challenge to ensure that all students have access to a quality education. However, some characterize the alternatives for rural families as being between schools that are too small or too far away, and for urban families as being between schools that are too large or too far away. State policies to equalize educational spending, cap bond and levy rates, require common teacher salary schedules, and consolidate school districts are geared toward improving quality and attempting to resolve problems of inefficiency. However, despite the use of state policy to equalize educational quality, educational gaps persist. The most identifiable achievement differences between and among school districts are well documented and result from variations in (a) teacher quality; (b) teacher mobility; (c) student characteristics, like poverty, ethnicity, gender, and mobility; and (d) community wealth (Goldhaber & Anthony, 2004; Haselton, 2004; Parcel & Dufur, 2001; Uribe, 2003, 2004).

Yu and Lau (2006) noted that no matter how well a systematic reform policy is conceptualized and planned, particular school district and school building constraints, resources, culture, leadership, and personnel particulars eventually determine whether the reform implementation will or will not be successful. The findings in this study support their conclusions that local contextual variables dramatically influence application of reform policy and thus student outcomes. While case study analyses have supported the importance of unique local variables on successful reform implementation, few quantitative or large-scale study methods have been explored to detect the presence or effects of these variations. Methods used in geospatial studies to detect such variation examine the concepts of spatial autocorrelation and spatial heterogeneity. The study reported in this article used these methods to analyze the relative student proficiency in school districts across two contiguous Midwestern states, Illinois and Iowa. The results support the use of more differentiated and district-level reform policy development rather than the centralized state and federal approaches that are dominant today.

Spatial Effects

Despite their prominence in economic research, there are very few educational studies (Rangel, 2006; Kilkenny & Haddad, 2008) that address issues related to spatial effects, specifically spatial autocorrelation and spatial heterogeneity (Anselin, 1988). *Spatial autocorrelation* occurs when value and location coincide (Anselin, 2001). In other words, spatial autocorrelation is when a value, like student test scores, in a location is correlated with its neighbors' values, rather than with another variable. *Spatial heterogeneity* exists when structural changes related to location are detected in the data set, implying unstable relationships between values of observations and detectable spatial regimes (Baumont, Ertur, & Le Gallo, 2003). To illustrate, student proficiency is often characterized by strong geographic patterns—particularly between rural and urban school districts—meaning that the spatial distribution of student proficiency is not stable across space. Previous studies on the effectiveness of educational reform indicate that policy applied on a grand systemic level may be ineffective because of local contextual variation (Rossman & Wilson, 1996; Conley & Picus, 2003). Previous studies have assumed that student testing trends within regions of a state and even across states have been the result of traditional variability in socioeconomic status, teacher quality, urban versus rural, district size, or level of parental education, to name a few variables (Baker, Keller-Wolff, & Wolf-Wendel, 2000; Stewart, 2008). In addition, purported achievement trends within many states have led to state-wide, one-size-fits-all policies, like district consolidation.

School District Consolidation

Pertinent for this study, consolidation proponents in Iowa cited lower student test scores among small, rural districts as grounds for school closure and district merger (Lambert, 2006). Linking low test scores and smaller district size is regularly used as evidence to support educational policy decisions nationally. In addition, consolidation policy is driven by claims of economic inefficiencies. In fact, some states (i.e., Maine, Oregon, Wyoming) have required all districts to close and merge if enrollment dips below a set target level. Consolidation advocates support these broad educational reform policies by suggesting that larger schools and districts can provide more options for students and attract more experienced teachers than small districts can.

In fact, the effects of district consolidation on student achievement are varied across the research literature. Consolidation proponents indicate a number of advantages including a broadened curriculum, greater funding and staffing flexibility, and increased capacity to support special needs and academically gifted students, thus providing a higher quality education both academically and socially (Nelson, 1985; Sell, Leistritz, & Thompson, 1996; Chance & Cummins, 1998). However, other research has questioned the merits of district consolidation noting that students in larger schools suffer from poor attendance, higher discipline incidents, more suspensions, and a decrease in personalized attention from teaching staff (Sher & Tompkins, 1977; Egelson, 1993; Reynolds, 2001).

Lee, Smerdon, Alfeld-Liro, and Brown (2000) noted that most student achievement variation resulted not from school or district size but from the quality of social relations within a school, mostly notably the supportive interactions between students and teachers. They concluded that since smaller districts and schools tend to have a higher frequency and continuity of social interactions among teachers and students, it might be surmised that small schools would promote achievement growth. Instead, Lee et al. (2000) found that although social interactions were more prevalent in small schools, the quality of interactions could either be positive or negative and thus help or hurt student performance depending on the specific school culture and individual teacher disposition. Therefore, although larger schools may display fewer positive social interactions, they may also encourage fewer negative social interactions that might lower student achievement.

Other studies concluded that power relations and school culture (Datnow, 2000), educator beliefs about students, and instructional quality (Datnow, Borman, Stringfield, Overman, & Castellano, 2003; Milesi & Gamoran, 2006) were the predominant variables impacting student performance rather than school or district size. However, since school and district size can influence variables like school culture, power relationships, and the number of social interactions, size remains a cogent factor in all of these studies. For example, Goddard (2003) pointed to the presence of increased social capital in larger districts and also concluded that higher social capital influences improved student achievement.

Despite the lack of consensus on the direct improvement of instructional quality as a result of consolidation, policymakers who support mergers generally focus on the purported economic efficiencies of consolidation and the link between increased per pupil expenditure and student test scores (Stremmick, 2001; Walberg & Fowler, 1987; Young, 1994) while others question the cost savings (Streifel, Foldsey, & Holman, 1991). In fact, Stiefel, Berne, Iatatola, and Norm (2000) found that student expenditures were not significantly different across very large and very small school districts. They concluded that given the similarity in expenditures, small schools should be favored because they help to promote success among minority and poor students, as some studies they cited suggest.

Consolidation and Spatial Effects

Although some have identified specific disadvantages of larger schools (DeYoung & Howley, 1992; Sher & Tompkins, 1977) and positive aspects of smaller schools (Fowler, 1992; Mullins, 1973), conflict continues between those advocating local control of rural schools and those who believe that educational quality and efficiency are improved in larger, more diverse districts (Chance & Cummins, 1998; DeYoung & Howley, 1992; Mullins, 1973).

Despite the debate, studies tend to concur that schools, especially in small rural districts and within certain geographic regions within a state, may be more heavily influenced by unique community social and cultural contexts that are difficult to identify and may have a significant impact on student performance. In fact, Alsbury and Shaw (2005), in a national study of district consolidations, revealed the importance and connectedness of community culture and schools. Similarly, Alsbury (2008) demonstrated a possible connection among dissatisfaction and controversy in small communities, school board forced turnover/defeat, and declining student achievement. These studies point to the importance of considering local community context as a variable in student achievement change. However, many existing studies do not incorporate a way to detect these types of regional data trends in their models. Spatial autocorrelation and spatial heterogeneity concepts can help assess such geographically-defined phenomena.

This methodological fault—not controlling for spatial effects—may lead to misspecifications of models, inefficient coefficients, and erroneous statistical inference (Anselin & Rey, 1991). If spatial effects are ignored, the results would be similar to the consequences of omitting a significant explanatory variable from the regression model, i.e., the estimators would be biased. Therefore, a need existed to control for spatial effects in a study comparing student proficiency change. By introducing the spatial dimension, this study overcomes the limitations of analyses that neglect to consider the dependency of each school district's geographic location on its quality of education. In regard to subsequent policy implications, if clusters of student proficiency within random state regions exist, after controlling for obvious, traditional demographic variance (socioeconomic status, teacher experience, etc.), it may follow that unique community social and cultural variables or other nontraditional spatial effects have a significant influence on educational quality. This would argue for a more differentiated approach to state-level policy development and raise questions as to the veracity of enacting state-wide policies, like district consolidation, that are structured around traditional demographic criteria like district enrollment.

Study Purpose

This study goes beyond the traditional boundaries of educational policy development and applies state-of-the-art spatial statistical and econometrics techniques to assess what influences educational quality as measured by student test scores. Specifically, we draw upon literature and spatial methods used in economics to evaluate, through a unique lens, trends in student achievement scores across the states of Illinois and Iowa. We believe that, when educational policymakers seek ways to improve the quality of education, the analytical techniques used in this study can assist in contextualizing and refining the policymaking process.

This study sought to answer two key questions. Do spatial effects matter for the distribution of eighth grade math and reading scores (MAT-8 and READ-8) in Illinois and Iowa? If spatial effects matter, is there a difference when spatially comparing MAT-8 and READ-8 scores and their correlates in Illinois and Iowa? The research hypotheses associated with the questions are: If the spatial distribution of test scores depicts spatial autocorrelation and spatial heterogeneity, then educational studies should incorporate location as a variable influencing student achievement trends. If states and regions within states present different results when controlling for spatial effects and demographic and organizational variables known to influence test score fluctuation, then educational policy should be differentiated between and within states.

Literature Review

The Influence of Size on Achievement

Empirical studies comparing the quality of education in school districts across states are scarce. Most published studies about the quality of education in the United States use data for a unique state, estimating cost functions or production functions, to examine some correlates of quality of education. In general, studies that compare states use aggregated data (Carnoy & Loeb, 2002) that may be masking internal differences. These studies attempt to relate the quality of education to a variety of traditional factors thought to influence student achievement, and they vary in their findings.

Heinesen (2005), in a Danish study, found when students attend primary education in school districts with a population of less than 15,000, they had lower probability of completing secondary education. Driscoll, Halcoussis, and Svorny (2003) presented a California school level study in which they examined the impact of school district size on academic outcomes. They noted that no other study included the effects of class size, school size, and district size together in one model. Elementary, middle, and high schools were analyzed separately, and controls were added to account for multiple student and teacher characteristics and environmental conditions (population density) that may affect achievement trends. They concluded that "students attending school in larger districts did not perform as well on standardized tests as those attending schools in smaller districts" (Driscoll et al., 2003, p. 200). Additionally, larger district size influenced student scores more negatively in elementary and middle schools, with the largest impact on middle school student performance.

In addition, Brasington (1997) examined whether a relationship between district enrollment and student performance existed in Ohio, and discussed the results as they related to district consolidation decisions. He concluded that "both [larger] school and district size seem related to a decrease in proficiency test performance" (p. 52). Thus, he concluded that district consolidation may negatively impact student performance. Similarly, when Stiefel et al. (2000) examined high school budgets and performance in New York City, they suggested "that the city might do well to continue to encourage the formation and continuing support of small high schools" (p. 37).

Lee and Loeb (2000) asserted that school size was linked to a significant difference in teacher attitudes in small versus large schools and affected student achievement both directly and indirectly. They used teacher and student surveys and hierarchical linear models to determine (a) if "school size was related to teachers' assessment of their colleagues' willingness to assume responsibility for students' academic and social development, once demographic characteristics of teachers and school are taken into account" (p. 8); (b) if "school collective responsibility" was related to student achievement (p. 9); and (c) if school size directly influenced student achievement, when adjusting analyses models to account for "school collective responsibility and student composition" (p. 9). They concluded:

> (S)mall schools are favored compared with medium-sized or larger schools. In small schools, teachers have a more positive attitude about their responsibility for students' learning and students learn more. Even after taking size into account, learning is also higher in schools with higher levels of collective responsibility. Thus, we conclude that school size influences student achievement directly and indirectly, through its effect on teachers' attitudes. (p. 3)

Andrews, Duncombe, and Yinger (2002) asked the question: Does empirical research on economies of size support state policies that encourage consolidation? Based on a literature review of studies elaborated during three decades (cost function and production function studies), the authors concluded that a potential instructional and administrative cost savings may exist by moving from a very small district (550 or fewer pupils) to a district with 2000-4000 pupils. However, they found findings from other studies were less consistent, indicating a negative influence of increased school enrollment and student achievement. They concluded "there is some evidence that moderately sized elementary schools (300-500 students) and high schools (600-900 students) optimally balance economies of size with the potential negative sides of large schools" (p. 245).

Very few studies apply spatial statistics and econometrics methods to issues related to education. Kilkenny and Haddad (2008) used the same methodology applied in this study to test the 'horizon effect' hypothesis, which states that if the performance of a group depends on the frequency of observations of maximum performances in a group, a larger size group will ultimately display higher overall performance. Their study uses data on the percentages of tenth grade students who passed the ninth grade math test across public schools in Ohio. They found there were positive and significant effects on student achievement based on increasing the number of grade levels in a school and increasing the size within each grade level. This was believed to be caused by the positive effects of students observing more high performing peers within their grade level, who possibly challenged their peer group to excel. However, these gains diminished as the overall school grew too large.

As we show in this brief literature review, spatial statistics and econometrics methods have been scarcely used to assess the quality of education. We also find mixed results in addressing questions of optimal district and school size. We believe our study comprehensibly addresses these methods, supporting their use in studies that incorporate geographic location to better guide policy developments like consolidation.

Theoretical Frameworks

This study focuses on the concepts of *spatial autocorrelation* and *spatial heterogeneity* (Anselin, 1988; Bailey & Gatrell, 1995; Rogerson, 2001), both popular in economic and geographic studies. Both of these concepts focus on the influence of location on values. The theories suggest, for example, that populations in close proximity influence characteristics, like student performance, even when factoring out predictable regional similarities that may affect achievement, like socioeconomic level, mobility rates, ethnicity rates, and others. Because of their absence in educational policy research studies, a further explanation of these concepts is warranted.

Spatial autocorrelation. Spatial autocorrelation can be positive when spatial clustering of high-high and low-low values are observed, or negative when spatial outliers of high-low and low-high values are observed (Guillain, Le Gallo, & Boiteux-Orain, 2006). For example, in the education context, a positive spatial autocorrelation may be present if school districts with high or low proficiency on standardized test scores are surrounded by neighboring school districts with similarly high or low scores, creating patterns of clustering within a region or state. Conversely, there may be a negative spatial autocorrelation if school districts with high or low proficiency on test scores are surrounded by districts with low or high proficiency, leading to the absence of patterns of clustering within a region or state. In this last case, locations tend to be surrounded by neighbors with very dissimilar values.

Spatial heterogeneity. Spatial heterogeneity exists when structural changes related to location are detected in the data set. In such cases, spatial regimes might be present, for example, when the data variations follow specific geographical patterns such as east, west, north, or south. When a variable is characterized by a distinct spatial distribution in different geographic sub-regions, these sub-regions might point to the existence of spatial regimes (Baumont et al., 2003). For example, school districts located in the north of a study area may be clustered around low test scores, while school districts located in the south of a study area may be clustered around high test values. In this example, north and south are known as the spatial regimes characterizing the spatial distribution of the student test scores. Conversely, a negative spatial autocorrelation is the absence of such clusters.

School reform approaches. In addition to the concepts of spatial autocorrelation and heterogeneity, two theoretical frameworks toward school reform efforts were applied: common and differentiated (Rumbarger & Palardy, 2005). The *common* framework suggests that a single approach or a single set of policy reform interventions can elicit change in student achievement at a larger regional, state, or national level regardless of unique school contextual features. Conversely, the *differentiated* framework supports contextualizing policy development and reform efforts at a more localized level.

This study attempts to measure if differences in scores still exist when controlling for spatial effects among districts in a state. This would support the individual identification of unique, regional contextual variables, and policy remedies utilizing a more localized, differentiated framework rather than the typical common framework of state-wide, one-size-fits-all reform policies that are widespread today. This calls for several basic questions. Can we expect a common national reform policy to be equally effective in Illinois and Iowa school districts? Do seemingly benign spatial variations—such as neighbors' achievement (spatial autocorrelation) or geographical location within the state (spatial heterogeneity)—influence student achievement?

Figure 1 displays our research framework in which the theoretical elements presented above are connected with all the variables used in the modeling. Details about the variables and the model are described below. It is important to highlight that a feedback loop is proposed so policymakers can revisit the model's output from time to time to ensure that their policies are doing the job they expect them to do.

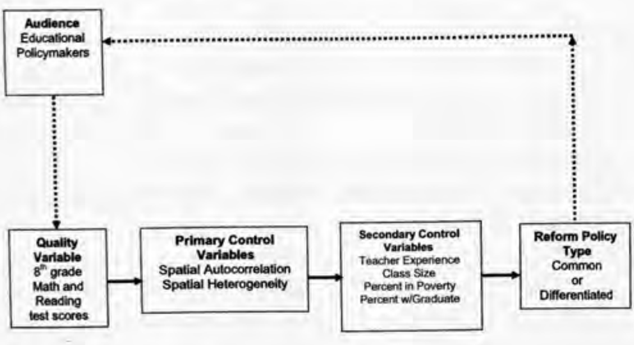

Figure 1. Research framework.

Study Methods

Data

In this study, we employed two dependent variables: percents of eighth grade students proficient in reading and proficient in math. In addition, we controlled for a number of variables that numerous research studies have concluded have a significant impact on student test scores. For example, to account for variations in teacher quality, we controlled for the average years of teacher experience. In the area of student/family demographics, we controlled for percent of school district population in poverty, and the educational level of the community as measured by percent of school district population that have a master's or Ph.D. degree. Our final control variable—related to school district structure—was class size measured by the pupil-to-teacher ratio. Initially, we intended to have other independent variables in our model specification such as total expenditure per pupil, a real average teacher salary, and the number of students in the whole school district. However, these variables were not significant when included in the specification. In addition, the model specification we adopted was the most parsimonious one, not performing any worse than the full specification in terms of goodness of fit. Table 1 displays the summary statistics of all variables used in the study.

Table 1

Summary Statistics of All Variables

	M		Minimum		Maximum		*SD*	
	IL	IA	IL	IA	IL	IA	IL	IA
Class size (pupils per teacher)	17	14	5	6	27	22	3	2
Teacher experience (years)	15	12	5	6	24	23	3	3
Reading proficiency (%)	71.01	70.07	21.10	42.86	100.00	91.49	12.69	8.00
Math proficiency (%)	57.33	73.35	4.90	46.51	100.00	91.95	17.42	8.61
% w/ graduate degree	0.0005	0.0314	0.0000	0.0029	0.3210	0.2612	0.0465	0.0254
% in poverty	0.0799	0.0830	0.0006	0.0207	0.4886	0.2254	0.0560	0.0321

These data came from two different sources. The dependent variables—teacher quality and school district structure—for the year 2002 came from the Illinois State Board of Education and the Iowa Department of Education and were collected at the school district level. So, the spatial unit of analysis is the school district. In 2002, Illinois had 893 school districts and Iowa had 370 school districts. Student/family demographics were collected at the U.S. Census Bureau webpage, for the year 2000.

Spatial Weight Matrix

To conduct the spatial analysis proposed in this study, it was necessary to define a spatial weight matrix W. This matrix imposes a neighborhood structure in the data and can be defined in a variety of ways. In other words, the matrix help us to understand which school district is neighbor to which. By examining empirical work and focusing on the type of spatial weight matrices used, the context they were used in, and their applicability to our study, we opted for one simple binary queen contiguity and two *k*-nearest-neighbors matrices per state. We used three matrices per state to test the robustness of our results.

The simple binary queen contiguity matrix is composed of 0 and 1: If school district i has a common boundary and/or vertex with school district j, then they are neighbors and $w_{ij} = 1$; if school district i does not have a common boundary and/or vertex with school district j, then they are not neighbors and $w_{ij} = 0$. The diagonal elements are set to 0 because they compare each district with itself. The *k*-nearest-neighbors weight matrix is defined as:

$$\begin{cases} w^*_{ij}(k) = 0 \quad if \quad i = j \\ w^*_{ij}(k) = 1 \quad if \quad d_{ij} \leq d_i(k) \quad and \quad w_{ij}(k) = w^*_{ij}(k) / \sum_j w^*_{ij}(k) \\ w^*_{ij}(k) = 0 \quad if \quad d_{ij} > d_i(k) \end{cases} \tag{1}$$

where $d_i(k)$ is a critical cut-off distance defined for each school district i; $d_i(k)$ is the k^{th} order smallest distance between school districts i and j such that each school district i has exactly k neighbors. For this study, $k = 5$ and $k = 6$ were applied for both states. These values are chosen because they represent the highest frequency in the distribution of connection between Illinois and Iowa school districts, based on the examination of the simple binary queen contiguity matrix; that is, the majority of Illinois and Iowa school districts have five or six neighbors (for Illinois 163 and 182 school districts, respectively; for Iowa 93 and 102 school districts, respectively). All matrices, the simple binary queen contiguity and two *k*-nearest-neighbors matrices, are row standardized so that each row sums up to 1.

Spatial Autocorrelation

To better understand the quality of education in Illinois and Iowa, we first introduced primary control variables, spatial dimension, into our analyses. We found that spatial autocorrelation and spatial heterogeneity characterize the distribution of MAT-8 and READ-8 in both states. Spatial autocorrelation can be measured by different statistics, one of them is the Moran's *I*. This statistic gives a formal indication to determine the extent of linear association between the values in a given location with values of the same variable in neighboring locations (Ertur & Le Gallo, 2003). Table 2 shows the Moran's *I* statistics for MAT-8 and READ-8—based on 999 permutations—for both states, using the three spatial weight matrices. The results indicate a positive spatial autocorrelation in the distribution of both variables in both states since the statistics were significant with p value = 0.001. These results suggest that distributions of the 2002 MAT-8 and READ-8 were by nature clustered. In other words, school districts with high (low) percentage proficient were localized close to school districts with high (low) percentage proficient. We can note that the results for Illinois present higher values than Iowa, implying an even more prominent clustering trend in Illinois.

Table 2

Moran's I Statistics for MAT-8 and READ-8

Spatial weight matrix	Illinois		Iowa	
	Read	Math	Read	Math
5 neighbors	0.2508	0.3258	0.1497	0.1677
6 neighbors	0.2358	0.3091	0.1316	0.1617
Queen	0.2454	0.3143	0.1670	0.1761

Spatial Heterogeneity

Concerning spatial heterogeneity, we used the Getis-Ord statistics, which allowed us to work with the entire sample. They allow detecting the presence of local spatial autocorrelation and determining spatial regimes for the school districts in Illinois and Iowa. Getis-Ord statistics were calculated for each school district for both variables and both states using the three spatial weight matrices (Le Gallo & Dall'erba, 2006). A positive value of this statistic for a school district indicates a spatial cluster of high proficiency values while a negative value of this statistic for a school district indicates a spatial cluster of low proficiency values. Based on these values, we determined our spatial regimes. This allowed us to develop a color-coded map showing MAT-8 and READ-8 spatial regimes within Illinois and Iowa, depicted in Figure 2. Generally, spatial regimes were persistent, highlighting some form of spatial heterogeneity.

Figure 2 illustrates the spatial regimes persistent for the three weight matrices. Generally low performing school districts tended to be clustered in certain parts of each state and high performers in different parts. When observing these maps, it is not straight forward to name these regimes in terms of the geography of the study areas—e.g. core versus periphery, or north versus south—since the distributions are not so homogenous. For example, the north of Illinois is characterized by high proficiency but the City of Chicago is low. Using a more generalized approach, we can state that in both—Illinois and Iowa—the north-south framework was persistent for both MAT-8 and READ-8. Low proficiency was located more in the south of the states and high proficiency more in the north, which highlights some form of spatial heterogeneity.

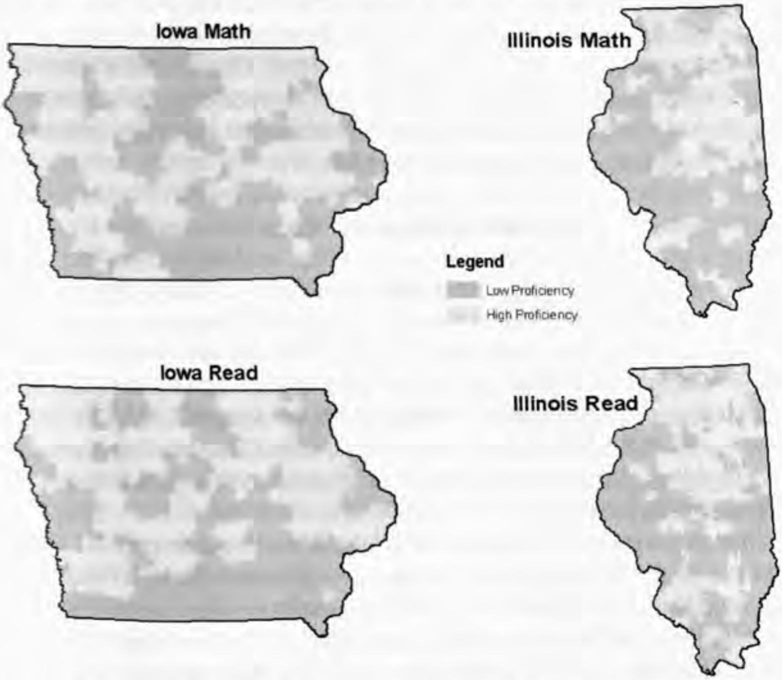

Figure 2. Spatial regimes based on Getis-Ord statistics.

Models

Based on the results presented above, we can state that location matters in the spatial distribution of both dependent variables in Illinois and Iowa. Therefore, we specified spatially-explicit regression models that incorporated spatial autocorrelation and spatial heterogeneity. As Le Gallo, Ertur, and Baumont (2003) state, "econometric estimations based on geographical data have to take into account the fact that [when a phenomena] spatially autocorrelated...spatial dependence between observations leads to unreliable statistical inference based on Ordinary Least Squares (OLS)" (p. 99).

We proposed a regression model for each state to examine quality of education and its correlates. The study used the MAT-8 and READ-8 standardized achievement tests as the dependent variables for the model, so the model was calculated twice: once using MAT-8 as the dependent variable, and once using READ-8 as the dependent variable. The independent variables for the regression model were: teacher experience, class size, percent in poverty, and percent with graduate degree.

We used OLS, spatial lag, and spatial error regressions. In the spatial lag model, spatial autocorrelation is handled by the endogenous spatial lag variable WMAT-8. According to Messner and Anselin (2004), "the spatial lag model implies that the geographic clustering of [student proficiency scores] is due to influence of [student proficiency scores] in one place on [student proficiency scores] in another. This model is consistent with some kind of diffusion process" (p. 138). According to Messner and Anselin (2004), "a spatial error model indicates that clustering reflects the influence of unmeasured variables" (p. 138).

Results and Spatial Trends

In addition to being displayed below in numerical tables, the results of the study are shown as visual, color-coded maps to better detect spatial trends throughout Iowa and Illinois. We reported results from the six-nearest neighbors weight matrix for both states. The estimations using the simple binary queen contiguity and the five-nearest neighbors weight matrices led to the same results and were, therefore, not reported here. This fact highlights the robustness of our results with regard to the choice of the spatial matrices.

It is important to emphasize that the R^2 criterion cannot be used to compare the "fit" of non-spatial OLS and spatial models, so we employed the Log Likelihood (i.e., the value of the maximum likelihood function), Akaike Information Criterion (AIC) (Akaike, 1974), and Schwartz Criterion (SC) (Schwarz, 1978). The best models are the ones resulting in the lowest values for these three statistics. In our case, the differences were too slight but showed that the spatial models (i.e., spatial lag and spatial error models) perform better than the non-spatial models (i.e., OLS) for all models presented here. These results indicated that the best models for Illinois and Iowa were indeed the spatial ones.

For Illinois, the model using MAT-8 as the dependent variable suggested a spatial lag model. The model using READ-8 as the dependent variable suggested a spatial error model, instead.[1] In contrast, for Iowan models—using MAT-8 and READ-8 as the dependent variables—there was strong evidence for the need for spatial error specification.[2]

Performance Trends

As shown in Figure 2, generally low performing school districts tended to be clustered in certain parts of each state and high performers in different parts. This seems unremarkable on the surface, until you realize that traditional variables, while explaining the location of some of the clusters of high or low scores, do not seem to clearly explain all of these cluster locations. For example, south-central and north-central Iowa show clusters of low performing districts as would be predicted. However, there are low-performing clusters in the similarly rural and impoverished northeast and northwest parts of the state. Likewise, the area in Illinois representing the suburbs of Chicago, a wealthy area, has clusters of high-performing districts; however, high performance clusters also occur in the west-central part of the state in equally poor areas. Also, in the poor, rural southern part of Illinois there were both clusters of higher than predicted and lower than predicted scores within the same demographic type and the same geographic area. Thus, low performing district clusters were not only located in high poverty, high minority regions, and high performers were not always found in wealthy areas. Some poor, rural areas with inexperienced teachers, low community education rates, and other traits negatively influencing student test scores seem to have not only a single high performing district but clusters of them.

Illinois and Iowa Models

Table 3 displays the results for the Illinois models.[3] All coefficients were significant and have the expected signs for all models. The results confirm findings in previous studies showing a link between each of these secondary control variables and student test results. Correlation analyses of teacher experience show that high student performance coincided with more years of teacher experience. Class size was negative, as expected, indicating that decreasing class size was associated with higher test scores. This finding concurs with findings by Pong and Pallas (2001), who compared eighth grade math achievement and class size in the U.S. and other countries. They stated that the U.S. "is the only country in which there is a negative association between math achievement and class size, after controlling for the effects of teacher, school, and classroom variables… provid[ing] tentative support for using small classes for U.S. eighth graders" (p. 269). Figures for percent in poverty tell us that districts with higher numbers of students living in low-income families were associated with lower student performances. Finally, analyses of percent with graduate degree indicated that student performance in school districts with higher numbers of individuals with master's and Ph.D. degrees is on average higher than in communities with lower educational levels.

Table 3

2002 Estimation Results for Illinois School Districts

Estimation IL Read	Non-spatial OLS		Spatial error ML		Groupwise ML	
β_0 (constant)	79.46	(0.000)	79.85	(0.000)	79.85	(0.000)
β_1 (teacher experience)	0.453	(0.000)	0.347	(0.000)	0.347	(0.011)
β_2 (class size)	-0.651	(0.000)	-0.592	(0.000)	-0.592	(0.000)
β_3 (percent w/ graduate)	82.72	(0.000)	84.09	(0.000)	84.09	(0.000)
β_4 (percent in poverty)	-104.28	(0.000)	-103.27	(0.000)	-103.27	(0.000)
R^2-adjusted	0.386		—		—	
Lambda (Λ)	—		0.205	(0.000)	0.208	(0.000)
AIC	5902.75		5886.26		5879.33	
SC	5926.13		5909.64		5902.71	
Log likelihood	-2946.38		-2938.13		-2934.66	
Likelihood ratio test	—		—		6.931	(0.000)
LM LAG	—		0.073	(0.786)	—	
Variance low proficiency	—		—		110.93	(0.000)
Variance high proficiency	—		—		83.36	(0.000)
Estimation IL Math	**Non-spatial OLS**		**Spatial lag ML**		**Groupwise ML**	
β_0 (constant)	63.99	(0.000)	46.11	(0.000)	47.86	(0.000)
β_1 (teacher experience)	1.039	(0.000)	0.99	(0.000)	0.943	(0.000)
β_2 (class size)	-0.958	(0.000)	-0.807	(0.000)	-0.824	(0.000)
β_3 (percent w/ graduate)	119.55	(0.000)	98.35	(0.000)	94.04	(0.000)
β_4 (percent in poverty)	-147.87	(0.000)	-129.98	(0.000)	-131.92	(0.000)
R^2-adjusted	0.415		—		—	
Lag	—		0.273	(0.000)	0.272	(0.000)
AIC	6365.5		6329.72		6326.18	
SC	6388.88		6357.78		6354.24	
Log likelihood	-3177.75		-3158.86		-3157.09	
Likelihood ratio test	—		—		3.538	(0.059)
LM ERROR	—		1.198	(0.273)	—	
Variance low proficiency	—		—		184.45	(0.000)
Variance high proficiency	—		—		149.65	(0.000)

From Table 3 we concluded that the best model for Illinois MAT-8 was the spatial lag incorporating groupwise heteroskedasticity; and the best model for Illinois READ-8 was the spatial error incorporating groupwise heteroskedasticity.

Table 4 displays the results for the Iowa models.[4] All coefficients were significant and had the expected signs for all models, except class size for the READ-8 model, which was not significant for all models. We concluded that the best model for Iowa MAT-8 was the spatial error, and the best model for Iowa READ-8 was the spatial error incorporating groupwise heteroskedasticity.

Table 4

2002 Estimation Results for Iowa School Districts

Estimation IA Read	Non-spatial OLS		Spatial error ML		Groupwise ML	
β_0 (constant)	68.81	(0.000)	68.68	(0.000)	68.52	(0.000)
β_1 (teacher experience)	0.476	(0.002)	0.39	(0.013)	0.393	(0.012)
β_2 (class size)	-0.126	(0.468)	-0.0419	(0.806)	-0.04	(0.811)
β_3 (percent w/ graduate)	91.1	(0.000)	90.7	(0.000)	90.8	(0.000)
β_4 (percent in poverty)	-66.54	(0.000)	-66.01	(0.000)	-65.8	(0.001)
R^2-adjusted	0.162		—		—	
Lambda (Λ)	—		0.043	(0.001)	0.043	(0.001)
AIC	2531.87		2519.05		2519.01	
SC	2551.43		2538.62		2538.58	
Log likelihood	-1260.93		-1254.53		-1254.50	
Likelihood ratio test	—		—		0.042	(0.978)
LM LAG	—		0.049	(0.944)	—	
Variance low proficiency	—		—		50.51	(0.000)
Variance high proficiency	—		—		52.71	(0.000)
Estimation IA Math	**Non-spatial OLS**		**Spatial error ML**			
β_0 (constant)	76.74	(0.000)	76.48	(0.000)		
β_1 (teacher experience)	0.516	(0.002)	0.349	(0.038)		
β_2 (class size)	-0.483	(0.010)	-0.317	(0.082)		
β_3 (percent w/ graduate)	104.12	(0.000)	108.14	(0.000)		
β_4 (percent in poverty)	-73.32	(0.000)	-75.309	(0.000)		
R^2-adjusted	0.156		—			
Lambda (Λ)	—		0.054	(0.000)		
AIC	2589.19		2568.37			
SC	2608.75		2587.94			
Log likelihood	-1289.59		-1279.18			
Likelihood ratio test	—		—			
LM LAG	—		0.875	(0.349)		
Variance low proficiency	—		—			
Variance high proficiency	—		—			

Residual Maps

In addition to the numerical representation of data, residual maps were elaborated based on the best model specification for each state MAT-8 model and READ-8 model. These maps depict the difference—for each school district—between the actual and fitted values of observations on the dependent variable and are a useful tool for visual inspection of patterns. We chose to use the standard deviation classification to elaborate these maps. These maps indicate tendencies of over-prediction (negative residuals) and under-prediction (positive residuals). They also help us to visualize the very large residuals, which are higher than 1.5 standard deviations, and lower than -1.5 standard deviations. These very large residuals are considered outliers, and indicate the existence of other variables where the independent variables and location were not sufficient to explain the dependent variable. If policymakers are interested in improving student proficiency, they should give special attention to these outliers because they represent a greater mismatch between predicted value and the actual student proficiency.

Figure 3 displays the residual maps for Illinois when predicting student proficiency using primary and secondary control variables. Notable trends include (a) dissimilar pattern of colored areas and location, i.e., similar colored areas did not tend to be in similar locations (random pattern); (b) very large residuals (darker colors) were randomly spread across the state's school districts; (c) urban areas did not perform as outliers; (d) rural areas did perform as outliers; (e) Chicago metropolitan area presented under-prediction trends (brownish colors); and (f) the City of Chicago was not considered a very large residual.

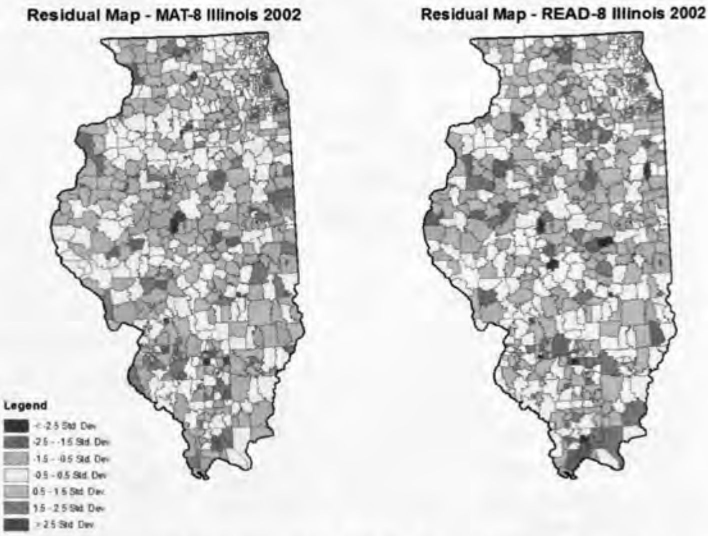

Figure 3. Illinois residual maps for MAT-8 and READ-8 models.

Table 5 depicts the distribution of under-predicted and over-predicted school districts for Illinois models. This table helps us to understand the presence of very large residuals. We are interested in paying more attention to 'extremely high' and 'very high' categories, since these categories indicate the existence of other variables where our primary and secondary control variables were not sufficient to explain student proficiency. Model MAT-8 had 12.9 percent of outliers, and model READ-8 12.3 percent of outliers. From the policymaker perspective, these school districts—considered outliers—should receive special attention in order to identify the other variables that are influencing student proficiency in their territory. It is important to highlight that they are in rural areas, not in urban areas.

Table 5

Percentage of Illinois School Districts Under- and Over-Predicted Based on the Residual Values

Illinois	Interval	Scale	Percentage
MAT-8	< -2.5 *SD*	Extremely high over-prediction	1.1%
	-2.5 – -1.5 *SD*	Very high over-prediction	5.7%
	-1.5 – -0.5 *SD*	High over-prediction	21.4%
	-0.5 – 0.5 *SD*	Over-prediction and under-prediction	42.8%
	0.5 – 1.5 *SD*	High under-prediction	22.9%
	1.5 – 2.5 *SD*	Very high under-prediction	5.4%
	> 2.5 *SD*	Extremely high under-prediction	0.7%
READ-8	< -2.5 *SD*	Extremely high over-prediction	1.3%
	-2.5 – -1.5 *SD*	Very high over-prediction	5.9%
	-1.5 – -0.5 *SD*	High over-prediction	18.9%
	-0.5 – 0.5 *SD*	Over-prediction and under-prediction	44.0%
	0.5 – 1.5 *SD*	High under-prediction	24.8%
	1.5 – 2.5 *SD*	Very high under-prediction	5.1%

Figure 4 displays the residual maps for Iowa when predicting student proficiency using primary and secondary control variables. Notable trends include (a) nonrandom patterns of colored areas and location, i.e., similar colored areas tended to occur in similar locations (presence of spatial autocorrelation); (b) very large residuals (darker colors) were randomly spread across the state's school districts; (c) urban areas that did not perform as outliers; and (d) rural areas that did perform as outliers.

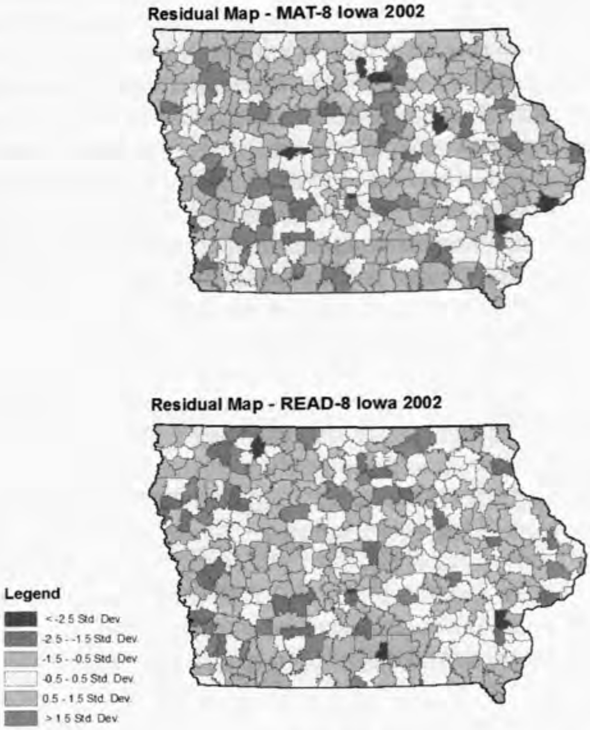

Figure 4. Iowa residual maps for MAT-8 and READ-8 models.

Table 6 depicts the distribution of under-predicted and over-predicted school districts for Iowa models. This table helps us to understand the presence of very large residuals. We are interested in paying more attention to 'extremely high' and 'very high' categories, since these categories indicate the existence of other variables where our primary and secondary control variables were not sufficient to explain student proficiency. Model MAT-8 contained 12.7% outliers, and model READ-8 contained 17.0% outliers. From the policymaker perspective, these school districts—considered outliers—should receive special attention in order to identify the other variables that are influencing student proficiency in their territory. It is important to highlight that they represent rural areas, not urban areas.

Table 6

Percentage of Iowa School Districts Under- and Over-Predicted Based on the Residual Values

Iowa	Interval	Scale	Percentage
MAT-8	< -2.5 *SD*	Extremely high over-prediction	1.9%
	-2.5 – -1.5 *SD*	Very high over-prediction	5.4%
	-1.5 – -0.5 *SD*	High over-prediction	21.6%
	-0.5 – 0.5 *SD*	Over-prediction and under-prediction	37.3%
	0.5 – 1.5 *SD*	High under-prediction	28.4%
	1.5 – 2.5 *SD*	Very high under-prediction	5.4%
READ-8	< -2.5 *SD*	Extremely high over-prediction	4.0%
	-2.5 – -1.5 *SD*	Very high over-prediction	7.3%
	-1.5 – -0.5 *SD*	High over-prediction	20.5%
	-0.5 – 0.5 *SD*	Over-prediction and under-prediction	39.5%
	0.5 – 1.5 *SD*	High under-prediction	25.9%
	1.5 – 2.5 *SD*	Very high under-prediction	5.7%

Conclusions and Implications

Nationally, educational practitioners and researchers have grappled with the desire to determine what causes school districts to perform at varying levels of achievement. Throughout the 1900s individual school districts were generally allowed to develop their own reform policies based on local context and need. In the late 1900s and early 2000s, a national cry for school accountability led to federal and state accountability measures that have driven a centralized policy development trend. Some opponents of common centralized policy have argued that local or regional contextual variables within and between states support the use of more differentiated and district-level policy development. However, empirical evidence to support these concerns has been scarce.

Through the use of spatial effects methods absent in other studies of this type, this study provides results that support the notion of differentiated policy development. In addition, results indicate differences between and within states possibly complicating the effectiveness of common policy development at both the national and state levels. Moreover, study findings provide implications for the use of particular control variables, the use of spatial methods, and regime mapping techniques. Finally, the study points to the ability of using residual maps to show the areas or districts within a state that are performing above or below what is predicted when controlling for recognized demographic variables that influence student achievement. As can be seen in this study, comparing raw mean student test data without controlling for the spatial effects used in this study can lead to potentially erroneous conclusion concerning the success or failure of district reform efforts and, subsequently, result in improper state policy decisions. Residual maps, as used in this study, reveal unique patterns of district-level student achievement values in specific geographic locations that differ from simply comparing trends in raw test scores. This points to the need for incorporating spatial effects when comparing district performance, and highlights the inappropriateness of using a single measure, like district size, as a criterion for mergers in state-wide consolidation policies.

Cogent Variables

The results shown in the Tables 3 and 4 spatial models suggest that educational policymakers should pay particular attention to teacher experience and class size when trying to improve the quality of education in lagging school districts. Generally, this suggests that if policymakers want to ensure that all students, regardless of their geographic location within a state, have access to quality education, they should (a) allocate more experienced teachers to disadvantaged locations, and (b) reduce class size if schools are experiencing crowded classrooms. However, school districts in this study revealing very large residuals suggest a need to explore additional variables influencing student proficiency in their local region of the state.

National Policy Efficacy

The fact that differential findings emerged between two states, even when using identical variable analyses and controlling for common demographic variation, further rejects the practice of relying too heavily on a national common policy or tightly-defined criteria when attempting to lead reform in local school districts. The problem is that without the proper analysis of unique spatial trends within regions of a state, as shown in this study, it is easy to make judgments using variables that predict student achievement in some states or regions but not others. For example, our study found low achievement clusters not only in urban but also in rural areas in both states. Thus, the variable of rurality or district enrollment may be less critical as performance predictors than, for example, levels of teacher experience. This fact also brings into question state-level policies, like those for school consolidation, that measure potential district success in terms of enrollment exclusively. This study makes it abundantly clear that both states had clusters of small, rural districts that were performing well and clusters of larger districts performing poorly. Such generalized analysis in Iowa, indeed, led lawmakers in the 2000s to the conclusion that schools of a certain small size perform more poorly, subsequently driving a call for a state-level consolidation policy (Lambert, 2006). Our study results show that this conclusion may be faulty given the broad method of analysis, and we recommend the incorporation of spatial effects in the modeling process when deciphering testing trends and their causes.

Using Residual Maps

Another advantage of using spatial econometrics is that, after estimations, we can produce residual maps that can guide policymakers to the exact locations where nontraditional reform efforts may be most needed. With respect to both Illinois and Iowa, maps in Figures 3 and 4 show specific areas in each state with large under-predicted performance (colored darker green) and large over-predicted performance (colored darker brown). This finding has significant interest to practitioners and policymakers alike on a variety of levels. These maps allow state officials to identify districts in need of targeted assistance. Perhaps even more notable are the locations that scored lower than predicted given their particular demographics. These districts are performing lower than should be expected and thus other typically unmeasured variables must be influencing their performance. State and local officials would be correct in calling for a thorough evaluation of these school districts and their communities to determine what localized external or internal issues were leading to performance problems. The identification of these needs could then be addressed using localized, differentiated policy and assistance.

From a policymaker's perspective, these darker areas are the ones that deserve first priority, if the goal is to improve quality of education. Resources could be targeted to those areas. Additionally, regionalized policy or more flexibility in state common policy could be employed to speak to these locations within the state. For example, it is interesting to observe that the City of Chicago stands out in relation to its more highly proficient neighbors in asking for and receiving state assistance (Illinois State Board of Education, 2007). Even foundations supporting educational reform have become nearly exclusive in targeting urban, high minority areas for assistance. However, rural, southern Illinois also has many lagging districts, and some might argue even less community social and cultural capacity to effect change. In Iowa, as noted earlier, locations that need scrutiny were found in rural areas.

These low performing outliers also provide us with another equally important piece of data. Specifically, some low achieving districts were predicted to be low achieving. These districts have predictable external challenges that make student performance gains much more difficult. While this does not excuse district educators from working toward program improvements and the success for every child, it does place them in a very different light in terms of reporting their results to parents and the community. In other words, poor performance in one community context should not be the same as poor performance in another. While not acceptable as an excuse for a lack of improvement focus, this understanding could affect the predisposition of the community members to educators, the message from state leaders to these districts, and, thus, the message of school leaders to teachers and students. This approach to improvement could allow for a more positive climate and supportive culture in the face of the expected educational challenge, rather than the presumption that staff performance is comparably poor to other districts across the state. In addition, state policy and assistance to these communities probably should focus as much on improving social realities in the community through initiatives like school/agency partnerships, as opposed to solely targeting the district program, organizational quality, and personnel.

Finally, the residual mapping used in this study provides us with another group of district performance outliers—those that performed higher than expected. These clusters of districts could provide pocket examples of programs to be visited, studied, and exemplified. It is interesting to note that these high performing outliers represented a wide variation in size, community demographics, and regional location in each state. They also were not necessarily obtaining the highest raw test scores in the state, but were high performers given their contextual circumstances and challenges. It is likely that some high scoring districts do so not because of notable programs or personnel, but because of their lower level of external challenge. When these districts are held up as exemplars, often districts in more challenging circumstances correctly discount their successes as a function of preferable demographics. This explains one reason districts in difficult settings are not able to easily find helpful improvement ideas or district exemplars to emulate given their unique challenges. Identifying exemplar districts using spatial econometrics could make significant contributions to discovering often hidden exemplary practice in districts of all demographic types.

Hidden Causation

While the scope of this study did not include the qualitative, in-depth analysis of performance outliers, a study by Delagardelle and Alsbury (2007) demonstrated the potential presence of local contextual variables within each state region that may be difficult to identify and measure. In a study linking school board member beliefs and student achievement variance, they noted that board members in the southern lowest performing region in Iowa were unique in their shared belief that schools could not reach every student. While these variables may merely be incidental and anecdotal, it supports the importance of spatial analysis to focused regionalized needs assessment and targeted assistance as opposed to the current district-level or state-level approaches currently popular.

While this study was not designed to explain all the reasons for the presence of spatial autocorrelation and spatial heterogeneity, the need for studies to include the measure of spatial effects with traditional variable analysis seems clear. Viewed from a substantive policy perspective, national leaders need to consider state-to-state variance made clear through spatial analysis, and local and state educators and policymakers need to consider regionalized reform efforts. The one-size-fits-all reform efforts may not work. From a practical standpoint, spatial effects need to be considered when measuring school performance for inclusion on the No Child Left Behind (NCLB) Schools in Need of Improvement (SINA) list. In addition, state-wide professional development initiatives should be regionalized to local districts or clusters of districts that are over-predicted on the residual map, rather than promotion of state-wide training initiatives. Conversely, regionalized programs, targeted by performance clusters identified in the spatial maps, may still be effective and provide greater resource efficiencies than district-by-district reforms. Finally, the transferability of reform programs may be improved by evaluating them against spatial effects.

The state-to-state comparisons in this study provide a previously unidentified spatial clustering that should inform state education reform policy. Where spatial cluster trends are verified, qualitative study methods could then be employed to analyze state policy, tax structure, economics, funding of teachers, or other state policy level variations within and across states that might be causing the regional clustering.

End Notes

[1] The tests for Illinois MAT-8 follow: Lagrange Multiplier Test for an autoregressive spatial lag variable (46.504, p value 0.000) and Lagrange Multiplier Test for a spatial autocorrelation of errors (43.783 and p value 0.000). The tests for Illinois READ-8 follow: Lagrange Multiplier Test for an autoregressive spatial lag variable (9.648 and p value 0.0018) and Lagrange Multiplier Test for a spatial autocorrelation of errors (12.492 and p value 0.0004).

[2] The tests for Iowa MAT-8 follow: Lagrange Multiplier Test for an autoregressive spatial lag variable (12.657 and p value 0.0003) and Lagrange Multiplier Test for a spatial autocorrelation of errors (14.765 and p value 0.0001). The tests for Iowa READ-8 follow: Lagrange Multiplier Test for an autoregressive spatial lag variable (7.441 and p value 0.006) and Lagrange Multiplier Test for a spatial autocorrelation of errors (8.793 and p value 0.003).

[3] For all OLS models, the Jarque-Bera test (p value 0.0009 for MAT-8 and p value 0.0000 for READ-8) does reject the assumption of normality of errors, and therefore, when specifying the spatial models, the Instrumental Variables (IV) method was used in addition to Maximum Likelihood (ML) to estimate the spatial models and confirm that the ML-based estimates were reliable. We estimate with both methods—IV and ML—leading to the same results, and therefore the IV estimations are not displayed on the tables. For the OLS models, the White test (both MAT-8 and READ-8 p value 0.0000) does reject homoskedasticity, and therefore we explored groupwise heteroskedasticity possibly associated to structural instability across regimes.

[4] For all OLS models, the Jarque-Bera test (*p* value 0.003) for MAT-8 does reject the assumption of normality of errors, and therefore, when specifying the spatial models, the Instrumental Variables (IV) method was used in addition to Maximum Likelihood (ML) to estimate the spatial models and confirm that the ML-based estimates were reliable. We estimate with MAT-8 methods—IV and ML—leading to same results, and therefore the IV estimations are not displayed on the tables. The Jarque-Bera test (*p* value 0.169) for READ-8 does not reject the assumption of normality of errors. For the OLS models, the White test (*p* value 0.2141) for MAT-8 does not reject homoskedasticity. The White test (*p* value 0.076) for READ-8 does reject homoskedasticity at 5 percent, and therefore we explored groupwise heteroskedasticity possibly associated to structural instability across regimes.

Author Note

The authors would like to thank Maureen Kilkenny for her insights and suggestions in the beginning of the research project and the financial support from the Iowa State University Research Grant. We also thank Marissa Moore, Isha Kaphle, and Lan Wei for their research assistance.

References

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control, AC-19,* 716–304.

Alsbury, T. L. (2008). School board member and superintendent turnover and the influence on student achievement: An application of the Dissatisfaction Theory. *Leadership & Policy in Schools, 7*(2), 202–229.

Alsbury, T. L., & Shaw, N. (2005) Policy implications for social justice in school consolidation. *Leadership and Policy in Schools, 4*(2), 105–126.

Andrews, M., Duncombe, W., & Yinger, J. (2002). Revisiting economies of size in American education: Are we any closer to a consensus? *Economics of Education Review, 21,* 245–262.

Anselin, L. (1988). *Spatial econometrics: Methods and models.* Dordrecht, Netherlands: Kluwer Academic.

Anselin, L. (2001). Spatial econometrics. In B. Baltagi (Ed.), *Companion to econometrics* (pp. 310–330). Oxford, UK: Basil Blackwell.

Anselin, L., & Rey, S. (1991). Properties of tests for spatial dependence in linear regression models. *Geographical Analysis, 23*(2), 112–131.

Bailey, P. A., & Gatrell, A. C. (1995). *Interactive spatial data analysis.* New York: Prentice-Hall.

Baker, B. D., Keller-Wolff, C., & Wolf-Wendel, L. (2000). Two steps forward, one step back: Race/ethnicity and student achievement in education policy research. *Educational Policy, 14*(4), 511–529.

Baumont, C., Ertur, C., & Le Gallo, J. (2003). Spatial convergence clubs and the European regional growth process, 1980–1995. In B. Fingleton (Ed.), *European regional growth: Advances in spatial science* (pp. 131–158). Berlin: Springer-Verlag.

Brasington, D. M. (1997). School district consolidation, student performance, and housing values. *Regional Analysis & Policy, 27*(2), 43–54.

Carnoy, M., & Loeb, S. (2002). Does external accountability affect student outcome? A cross-state analysis. *Educational Evaluation and Policy Analysis, 24*(4), 305–331.

Chance, E., & Cummins, C. (1998). School community survival: Successful strategies used in rural school district consolidation. *Rural Educator, 20*(2), 1–7.

Conley, D. T., & Picus, L. O. (2003). Oregon's quality education model: Linking adequacy and outcomes. *Educational Policy, 17,* 586–612.

Datnow, A. (2000). Power and politics in the adoption of school reform models. *Educational Evaluation and Policy Analysis, 22,* 357–374.

Datnow, A., Borman, G. D., Stringfield, S., Overman, L. T., & Castellano, M. (2003). Comprehensive school reform in culturally and linguistically diverse contexts: Implementation and outcomes from a four-year study. *Educational Evaluation and Policy Analysis, 25,* 143–170.

Delagardelle, M. L., & Alsbury, T. L. (2007, April). *Roles and responsibilities of local school board members in relation to student achievement.* Paper presented at the annual meeting of the American Educational Research Association, Chicago, IL.

DeYoung, A. J., & Howley, C. B. (1992). The political economy of rural school consolidation. *Peabody Journal of Education, 67*(4), 63–69. (ERIC Document Reproduction Service No. ED 347018)

Driscoll, D., Halcoussis, D., & Svorny, S. (2003). School district size and student performance. *Economics of Education Review, 22,* 193–201.

Egelson, P. E. (1993, April). *Change for worse? One teacher's experience of rural school consolidation.* Paper presented at the annual meeting of the American Educational Research Association, Atlanta, GA.

Ertur, C., & Le Gallo, J. (2003). An exploratory spatial data analysis of European regional disparities, 1980–1995. In B. Fingleton (Ed.), *European regional growth: Advances in spatial science* (pp. 55–98). Berlin: Springer-Verlag.

Fowler, W. J. (1992, April). *What do we know about school size: What should we know?* Paper presented at the annual meeting of the American Educational Research Association, San Francisco, CA.

Goddard, R. D. (2003). Relational networks, social trust, and norms: A social capital perspective on students' chances of academic success. *Educational Evaluation and Policy Analysis, 25,* 59–74.

Goldhaber, D., & Anthony, E. (2004). *Can teacher quality be effectively assessed?* (Report #410958) Washington, DC: Urban Institute.

Guillain, R., Le Gallo, J., & Boiteux-Orain, C. (2006). Changes in spatial and sectoral patterns of employment in Ile-de-France, 1978–1997. *Urban Studies, 43*(11), 1–24.

Haselton, W. B. (2004). School finance reform in Kentucky: Impact on student outcomes. *Dissertation Abstracts International A65*(01), 1672. (UMI No. 3134186)

Heinesen, E. (2005). School district size and student educational attainment: Evidence from Denmark. *Economics of Education Review, 24*(6), 677–689.

Illinois State Board of Education. (2007). *Net disbursements/expenditure comparison, Title I-low income funding table by county school district.* Retrieved on October 10, 2007, from http://webrpt.isbe.net/webreports/WebReportsCR.aspx

Kilkenny, M., & Haddad, M. (2008). Rural human capital development. In J. Wu, P. W. Barkley, & B. A. Weber (Eds.), *Frontiers in resource and rural economics: Human-nature, rural-urban interdependencies* (Chapter 9). Washington, DC: Resources for the Future Press.

Lambert, R. (2006). Small districts: An achievement strategy for Iowa. *Rural Policy Matters, 8*(4), 4.

Lee, V. E., & Loeb, S. (2000). School size in Chicago elementary schools: Effects on teachers' attitudes and students' achievement. *American Educational Research Journal, 37*(1), 3–31.

Lee, V. E., Smerdon, B. A., Alfeld-Liro, C., & Brown, S. L. (2000). Inside large and small high schools: Curriculum and social relations. *Educational Evaluation and Policy Analysis, 22,* 147–171.

Le Gallo, J., & Dall'erba, S. (2006). Evaluating the temporal and spatial heterogeneity of the European convergence process, 1980–1999. *Journal of Regional Science, 46*(2), 269–288.

Le Gallo, J., Ertur, C., & Baumont, C. (2003). A spatial econometrics analysis of convergence across European regions, 1980–1995. In B. Fingleton (Ed.), *European regional growth: Advances in spatial science* (pp. 99–129). Berlin: Springer-Verlag.

Messner, S., & Anselin, L. (2004). Spatial analyses of homicides with areal data. In M. Goodchild & D. Janelle (Eds.), *Spatially integrated social science* (pp. 127–143). Oxford, UK: Oxford University Press.

Milesi, C., & Gamoran, A. (2006). Effects of class size and instruction on kindergarten achievement. *Educational Evaluation and Policy Analysis, 28,* 287–313.

Mullins, C. (1973). School district consolidation: Odds are 2 to 1 it'll get you. *American School Board Journal, 160*(11), 23–26 & 57.

Nelson, E. (1985). School consolidation. *ERIC Digest, 13,* 1–4.

Parcel, T. L., & Dufur, M. J. (2001). Capital at home and at school: Effects on student achievement. *Social forces, 79*(3), 881–912.

Pong, S., & Pallas, A. (2001). Class size and eighth-grade math achievement in the United States and abroad. *Educational Evaluation and Policy Analysis, 23*(3), 251–273.

Rangel, C. (2006). *How do schools and neighborhoods matter for student achievement? A study of educational inequality in Cartagena, Columbia.* Unpublished master's thesis, University of Illinois at Champaign-Urbana.

Reynolds, D. R. (2001, May). *Rural school consolidation in early twentieth century Iowa: Lessons for the early twenty-first century.* Paper presented at the annual meeting of the Wallace Family Conference on Gifted Education in Rural Schools, Iowa City, IA.

Rogerson, P. A. (2001). *Statistical methods for geography.* London: Sage.

Rossman, G. B., & Wilson, B. L. (1996). Context, courses, and the curriculum local responses to state policy reform. *Educational Policy, 10,* 399–421.

Rumbarger, R. W., & Palardy, G. J. (2005). Does segregation still matter? The impact of student composition on academic achievement in high school. *Teachers College Record, 107*(9), 1999–2045.

Schwartz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics, 6,* 461–464.

Sell, R. S., Leistritz, F. L., & Thompson, J. M. (1996). *Socio-economic impacts of school consolidation on host and vacated communities.* (Agricultural Report No. 347). Fargo, ND: North Dakota Agricultural Experiment Station. (ERIC Document Reproduction Service No. ED423100)

Sher, J. P., & Tompkins, R. B. (1977). Economy, efficiency and equity: The myths of rural school and district consolidation. In J. P. Sher (Ed.), *Education in rural America* (pp. 43–77). Boulder, CO: Westview Press.

Stewart, E. B. (2008). School structural characteristics, student effort, peer associations, and parental involvement: The influence of school- and individual-level factors on academic achievement. *Education and Urban Society, 40*(2), 179–204.

Stiefel, L., Berne, R., Iatarola, P., & Norm, F. (2000). High school size: Budgets and performance in New York City. *Educational Evaluation and Policy Analysis, 22*(1), 27–39.

Streifel, J. S., Foldsey, G., & Holman, D. M. (1991). The financial effects of consolidation. *Journal of Research in Rural Education, 7*(2), 13–20.

Stremmick, P. R. (2001). School district consolidation in Montana. *Dissertation Abstracts International, 52*(2), 469.

Uribe, C. (2003). Growing gains or growing pains? The potential impact of school expansion policies on student achievement, in Bogota, Columbia. *Dissertation Abstracts International, A64*(8), 2738. (UMI No. 3100173)

Uribe, S. (2004). An analysis of high and low wealth school districts' finance in Texas. *Dissertation Abstracts International, A65*(8), 2904. (UMI No. 3143603)

Wahlberg, H. J., & Fowler, W. J., Jr. (1987). Expenditure and size efficiencies of public school districts. *Educational Researcher, 16*(7), 5–13.

Young, E. (1994). Questioning consolidation. *Tennessee School Boards Journal, 11*(2), 33–37.

Yu, W., & Lau, C. (2006). Understanding education reform: Insights from stories of the changing school context. *Planning and Changing, 37*(3/4), 219–233.

Monica A. Haddad is an Assistant Professor and GIS Extension Specialist in the Department of Community and Regional Planning at Iowa State University, Ames, Iowa.

Thomas Alsbury is an Associate Professor in the Department of Educational Leadership and Policy Studies at North Carolina State University, Raleigh, North Carolina.